Management and spatial resolution effects on yield and water balance at regional scale in crop models

Julie Constantin, Helene Raynal, Eric Casellas, Holger Hoffman, Marco Bindi, Luca Doro, Henrik Eckersten, Thomas Gaiser, Balasz Grosz, Edwin Haas, et al.

Management and spatial resolution effects on yield and water balance at regional scale in crop models

Constantin Julie[1]; Raynal Helene[2]; Casellas Eric[2]; Hoffmann Holger[3]; Bindi Marco[4]; Doro Luca[5,6]; Eckersten Henrik[7]; Gaiser Thomas[3]; Grosz Balász[8]; Haas Edwin[9]; Kersebaum Kurt-Christian[10]; Klatt Steffen[9]; Kuhnert Matthias[11]; Lewan Elisabet[12]; Maharjan Ganga Ram[3]; Moriondo Marco[13]; Nendel Claas[10]; Roggero Pier Paolo[5]; Specka Xenia[10]; Trombi Giacomo[4]; Villa Ana[12]; Wang Enli[14]; Weihermüller Lutz[15]; Yeluripati Jagadeesh[11]; Zhao Zhigan[14]; Ewert Frank[3,10]; Bergez Jacques-Eric[1]

[1] AGIR, Université de Toulouse, INRA, Castanet-Tolosan, FR

[2] INRA U0875 MIAT, F-31326 Auzeville, FR

[3] Crop Science Group, INRES, University of Bonn, Katzenburgweg 5, 53115 Bonn, DE

[4] Department of Agri-food Production and Environmental Sciences, University of Florence, Florence, IT

[5] Desertification Research Centre, University of Sassari, Viale Italia 39, 07100 Sassari, IT

[6] Texas A&M AgriLife Research, Blackland Research and Extension Center, Temple, TX, USA

[7] Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 750 07 Uppsala, SE

[8] Thünen-Institute of Climate-Smart-Agriculture, Bundesallee 50, 38116 Braunschweig, DE

[9] Institute of Meteorology and Climate Research – Atmospheric Environmental Research, Karlsruhe Institute of Technology, Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, DE

[10] Leibniz Centre for Agricultural Landscape Research, ZALF, 15374 Müncheberg, DE

[11] Institute of Biological and Environmental Sciences, University of Aberdeen, 23 St Machar Drive, Aberdeen AB24 3UU, UK

[12] Department of Soil and Environment, Swedish University of Agricultural Sciences, Lennart Hjelms väg 9, 750 07 Uppsala, SE

[13] CNR-Ibimet, Florence, IT

[14] CSIRO Land and Water, Clunies Ross Street, Canberra, ACT, AU

[15] Institute of Bio- & Geosciences, Agrosphere (IBG-3), Forschungszentrum Jülich, 52425 Jülich, DE

Abstract

Due to the more frequent use of crop models at regional and national scale, the effects of spatial data input resolution have gained increased attention. However, little is known about the influence of variability in crop management on model outputs. A constant and uniform crop management is often considered over the simulated area and period. This study determines the influence of crop management adapted to climatic conditions and input data resolution on regional-scale outputs of crop models. For this purpose, winter wheat and maize were simulated over 30 years with spatially and temporally uniform management or adaptive management for North Rhine-Westphalia (~34 083 km²), Germany. Adaptive management to local climatic conditions was used for 1) sowing date, 2) N fertilization dates, 3) N amounts, and 4) crop cycle length. Therefore, the models were applied with four different management sets for each crop. Input data for climate, soil and management were selected at five resolutions, from 1×1 km to 100×100 km grid size. Overall, 11 crop models were used to predict regional mean crop yield, actual evapotranspiration, and drainage. Adaptive management had little effect (<10 % difference) on the 30-year mean of the three output variables for most models and did not depend on soil, climate, and management resolution. Nevertheless, the effect was substantial for certain models, up to 31 % on yield, 27 % on evapotranspiration, and 12 % on drainage compared to the uniform management reference. In general, effects were stronger on yield than on evapotranspiration and drainage, which had little sensitivity to changes in management. Scaling effects were generally lower than management effects on yield and evapotranspiration as opposed to drainage. Despite this trend, sensitivity to management and scaling varied greatly among the models. At the annual scale, effects were stronger in certain years, particularly the management effect on yield. These results imply that depending on the model, the representation of management should be carefully chosen, particularly when simulating yields and for predictions on annual scale.

Keywords: drainage, evapotranspiration, aggregation, decision rules, scaling

1. Introduction

Large-scale assessment studies based on simulations by crop models are frequently used to evaluate the impacts of agriculture. These studies usually focus on predictions of crop production in different contexts, such as climate change, its inter-annual variability, or trends over time (Gaiser et al., 2010; Nendel et al., 2013). Crop models are also used to study carbon sequestration or the greenhouse gas balance at regional or national scale (Gaiser et al., 2009, 2008; Tornquist et al., 2009). Other studies focus on the water balance and its dynamics at the watershed scale. For the latter, crop models are combined with other models (e.g., hydrological) and applied to quantitative water management and irrigation issues (Noory et al., 2011; Robert et al., 2018; Therond et al., 2014).

Crop models are useful tools for large-scale assessment since exhaustive measurements are not feasible or available. However, they were developed to simulate homogeneous fields, each represented by a combination of one soil and one climate. Some of these models were designed to simulate only one season, e.g. one crop and its management, while others are capable of simulating different crops in sequence, mimicking a crop rotation over a longer time period (Kollas et al., 2015). When applied at a larger scale, these models are usually applied in a gridded approach, simulating each grid cell independently, while assuming homogeneity within each grid cell (De Wit et al., 2012; Huang et al., 2015; Mo et al., 2005; van Ittersum et al., 2013). For such approach, it is necessary to provide input data for soil, climate, and management for each simulated unit. Depending on the study and the systems' heterogeneity, the number of homogeneous units can range from a few to millions. Such data, especially management data, are not easily available at large scales and at high spatial or temporal resolution. Several methods exist to scale-up the data over the whole study area, such as sampling, aggregation from fine to coarser resolution, extrapolation or interpolation of the available data (Ewert et al., 2011). As an alternative, management information can also be simulated for large-scale studies (Hutchings et al., 2012).

Nowadays, it is possible to obtain soil and climate data at a relatively high resolution and at a large or even global scale from databases such as those in the Global Soil Map project (http://globalsoilmap.net/), the European soil portal for soil, the SoilGrids project (soilgrids.org) and the international CORDEX initiative for climate projection (https://www.euro-cordex.net/). On the other hand, the available databases on crop management data are at coarser resolutions such as those reported by Portmann et al. (2010) and Sacks et al. (2010) for crop growing periods or earthstat.org for fertilizer inputs. Usually, the few data available on crop management come from interviews with farmers, local experts, or observation networks. It provides an average date of sowing, harvest, and fertilization for instance or fertilizer input amounts for a given region for different crops and generally concern only one or a few years. Some initiatives such as the observation network of the German weather service DWD documenting key phenological stages as well as sowing and harvest could provide useful data for regional modelling (Kersebaum and Nendel, 2014) but do not cover the wide range of cultivation operations such as nitrogen fertilization for instance. As a result, large-scale studies usually consider management as uniform across the region and fixed over multiple years. However, it is well known that crop management, such as sowing, varies over space and time (Leenhardt and Lemaire, 2002). Additionally, the sowing date significantly impacts crop development and yield (Bonelli et al., 2016), and influences subsequent management actions during season.

To address the scarcity of the data and to adapt the management to the local and annual conditions, some authors suggested using management rules. Such management rules aim at reproducing the behavior of farmers and their crop management strategies (Maton et al., 2005; Nendel, 2009; Senthilkumar et al., 2015). In addition, these rules would help identify better management strategies. For example, suitable climate and soil conditions could be identified to perform cultivation operations (e.g., avoiding soil compaction by triggering an operation when the soil is not too wet or avoiding the risk of frost for spring crops). This adaptive management, based on management decision rules, could have a strong impact on model outputs but is rarely investigated at a large scale. Since the impact of input data aggregation and adaptive management can differ according to the output variables and crop models, these effects should be investigated with respect to a range of different crop models, output variables, and cultivation operations (i.e. sowing, soil tillage, irrigation…).

The objective of this study was to analyze the effect of adaptive management and spatial resolution on regional yields, evapotranspiration, and drainage predicted by a set of crop models. The main issues addressed were (1) whether adaptive management and/or input resolution influence the crop models' outputs at the regional scale, in which way and how much and (2) whether the scaling effect varies when management changes over time and space.

To meet this goal, we quantified the impact of adaptive management and input resolution on the regional mean of simulated yield, evapotranspiration, and drainage for each individual year as well as for the 30-year average. We further analyzed whether the impact of management or spatial resolution depended on the crop model, output of interest, crop, or cultivation operation. To do so, we introduced adaptive management for sowing dates, fertilization dates, and crop maturity classes based on decision rules and variable amounts of nitrogen fertilization.

2. Materials and Methods

2.1. Study area

The study area was the 34.083 km^2 federal state of North Rhine-Westphalia (NRW, 6.0-9.5° E, 50.0-52.5° N), located in the west of Germany. NRW has a temperate and humid climate with an oceanic influence. Like Hoffmann et al. (2016b) and Zhao et al. (2015), we assumed in the simulations that agricultural land covered the entire region and that winter wheat and silage maize were the two dominant monoculture crops. Over the period studied (1982-2012), mean annual temperature was 9.7 °C, mean annual precipitation was 899 mm, and mean annual global radiation was 3.758 MJ m^{-2}.

2.2. Crop models

We selected 11 crop models to run the simulations from 1982-2012: AgroC (Herbst et al., 2008; Klosterhalfen et al., 2017), APSIM-Nwheat (Asseng et al., 2000), CoupModel (Conrad, 2009; Jansson, 2012), DailyDayCent (Del Grosso et al., 2006; Yeluripati et al., 2009), EPIC (Williams, 1995; Williams et al., 1983), Expert-N (Priesack et al., 2006), HERMES (Kersebaum, 2007), LINTUL in the framework solution SIMPLACE<Lintul5, SLIM> (Gaiser et al., 2013; Zhao et al., 2015b), MCWLA (Tao et al., 2009; Tao and Zhang, 2013), MONICA (Nendel et al., 2011) and STICS within the RECORD platform (Bergez et al., 2014; Brisson et al., 2003). These process-based models run at a daily time step, except for Expert-N, which runs at an hourly time step. The models represent soil and crop processes with differing degrees of simplification. All simulated winter wheat, but only seven simulated silage maize in this paper. All represent water and nitrogen stresses, except for AgroC and MCLWA, which represent only water stress.

2.3. Input data of the crop models

2.3.1. Climate and soil data aggregation

For climate, we used 30 years of daily weather for 34.168 grid cells of 1×1 km resolution and aggregated these data for the 10×10, 25×25, 50×50, and 100×100 km grid cells, as described by Hoffmann et al. (2015). For soil data, we used the dominant soil of the 1×1 km grid cells to set the soil type for the 10×10, 25×25, 50×50, and 100×100 km grid cells, respectively. For the soil and climatic data, see Hoffmann et al. (2016a) and for more details of data aggregation, see Hoffmann et al. (2016b). Figure 1 presents the maps of mean annual precipitation and available water capacity (soil water content at field capacity minus the soil water content at wilting point) of the soils for each resolution.

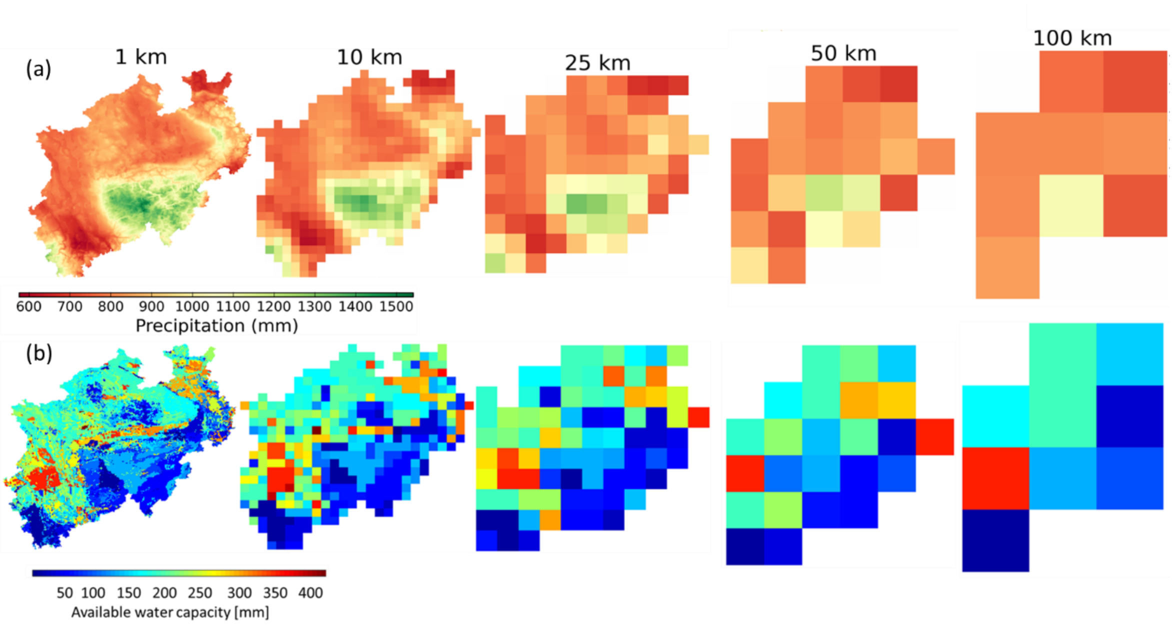

Figure 1. Maps of (a) mean annual precipitation over 30 years (1982-2012) and (b) available water capacity in each of the five resolutions for North Rhine-Westphalia, Germany.

All simulations were run using the same resolution of soil and climate data (km×km): 1×1, 10×10, 25×25, 50×50, and 100×100.

2.3.2. Crop choice and management sets

We simulated the two dominant crops of the region in monoculture in continuous model runs of 30 years on every grid cell. Both winter wheat and silage maize were grown under rainfed conditions and with mineral N fertilization (208 and 238 kg N ha^{-1} yr^{-1}, respectively). For both crops, we simulated export of crop residues at harvest and plowing of soil in autumn. We simulated six sets of management strategies to analyze the impact of adaptive management in interaction with the scaling effect:

1. M_{fix} is the reference, which is the same uniform management for a given crop regardless of the year or grid cell. We used the common cultivation operations in NRW as the reference management strategy. Winter wheat and silage maize were sown on 1st of October and 20th of April, respectively. Crops were

harvested at maturity or on 1st of August for wheat and 20th of September for maize, depending on the model.

2. *Ms* uses variable sowing, fertilization, and harvest dates for each cell, at each resolution and year according to decision rules based on climate, as in Senthilkumar et al. (2015) for maize and Savin et al. (2007) for wheat. For each crop, we calculated the earliest sowing date for all 30 years per grid cell (Fig. 2). Then, beginning on this date each year for each grid cell, we checked whether daily temperature and soil trafficability exceeded thresholds necessary for sowing. If all conditions were met, the crop was sown on that day. If sowing was impossible before a latest "allowed" date, it occurred on this date.

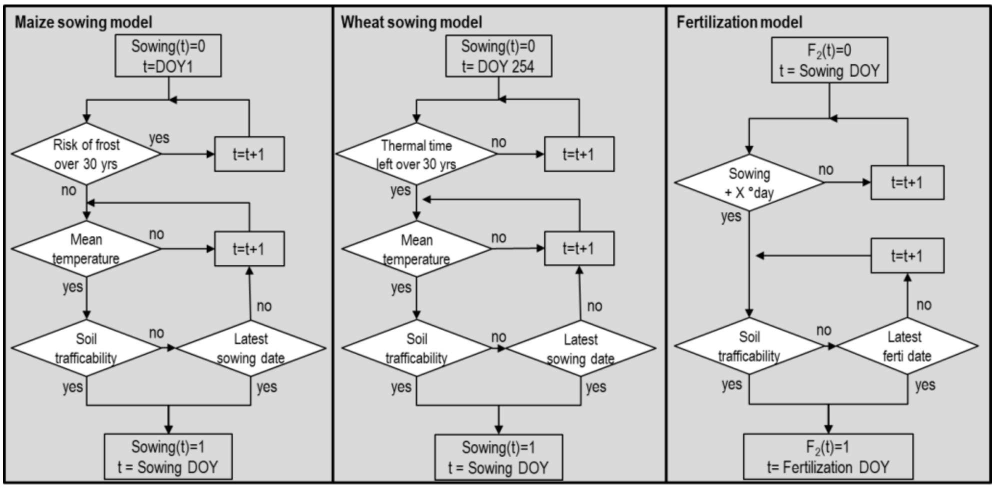

Figure 2. Overview of decision rules for wheat and maize sowing and fertilization dates. DOY = day of year

Fertilization date was set from the sowing date and depended on a minimum amount of thermal time and sufficient soil trafficability. Like for sowing, we defined a latest "allowed" date. We calculated the earliest harvest date as the number of days required to reach a certain cumulative thermal time from the sowing date. Beginning on this date, we checked soil trafficability each day to identify the first suitable harvest date. We calibrated the thresholds used in the decision rules to ensure that average dates were similar to those in M_{fix}. Estimated sowing dates among all grid cells and years ranged from 12th of March to 11th of May for maize and 21st of September to 16th of December for winter wheat. When averaged for all cells in the region, the mean sowing date each year ranged from 13th of April to 30th of April for maize and 22nd of September to 25th of October for wheat over the 30 years. Median sowing dates over the 30 years were 19th of April and 4th of October for maize and wheat, respectively, which were similar to those of M_{fix} (20th of April and 1st of October). Distributions of regional sowing dates for the five resolutions were similar, despite some differences for the coarser resolutions. Depending on the year, the mean regional sowing date was similar among resolutions.

3. Ms_{var} is similar to the *Ms* approach, but with the maturity class of the cultivar adapted to the climate conditions in each grid cell on each resolution. We chose one of three maturity classes or varieties (early, middle, or late) with a development length better adapted to climate characteristics by calculating the mean cumulative thermal time between sowing dates and the mean harvest date (20th of September for maize and 10th of July for wheat) over the 30 years. The maturity class in a given cell remained the same for all 30 years. We calibrated the three varieties for each model using the sowing and harvest dates of ten contrasting cells.
4. The fourth to sixth sets are the same as the *Ms* approach, but with a decrease in mineral N fertilization by 25% (Ms_{F75}), 50% (Ms_{F50}), and 75% (Ms_{F25}) of the reference fertilization amount, respectively. Thus, mineral N fertilization decreased from 238 to 179, 119 and 60 kg N ha^{-1} yr^{-1} for maize and from 208 to 156, 104 and 52 kg N ha^{-1} yr^{-1} for wheat in $Ms_{F75,}$ Ms_{F50}, and Ms_{F25}, respectively.

The objective of these six sets was to create spatial and temporal variability in the cultivation operations to analyze their impacts on the model results. The adaptive management based on climatic conditions was calculated for each grid cell for each of the five resolutions. The purpose was not to reproduce the actual management strategies, but to reproduce a credible range of cultivation operations over time and within the region to analyze their potential impacts on model outputs. Other cultivation operations such as tillage were assumed spatially and temporally uniform for all management sets.

2.3. Simulation overview and data selection

We analyzed three output variables: crop yield and two components of the water balance, evapotranspiration over the growing period and annual drainage under wheat to determine if some model outputs were more sensitive to scaling or management than others. Yield is often studied at large scale, while water fluxes are quite important when crop models are coupled with hydrological models to analyze water management at the watershed scale. We first selected and summarized simulated data (Table 1). We analyzed all three variables for five models only but yield and evapotranspiration were provided for six other models. Due to the complexity of the simulated experiments and model limitations, not all simulations were performed with all models (Table 1).

Table 1. Overview of the simulated resolutions and outputs analyzed by model, crop and management set.

			Resolution for Wheat						Resolution for Maize					
Model	**Code**	**Outputs**	**M_{fix}**[a]	**Ms**	**Ms_{var}**	**Ms_{F75}**	**Ms_{F50}**	**Ms_{F25}**	**M_{fix}**	**Ms**	**Ms_{var}**	**Ms_{F75}**	**Ms_{F50}**	**Ms_{F25}**
MONICA	MONI	Y, E, D[b]	All[c]	All	All	All	All	All	All	All	All	All	All	All
STICS	STIC	Y, E, D	All	All	All	All	All	All	All	All	All	All	All	All
LINTUL	LINT	Y, E, D	All	All	-	All	All	All	All	All	All	All	All	All
CoupModel	COUP	Y, E, D	All	All	-	All	All	All	-	-	-	-	-	-
Expert-N	EXPN	Y, E, D	-	All	All	All	All	All	-	-	-	-	-	-
EPIC	EPIC	Y, E	All	All	All	All	All	All	All	All	All	All	All	All
HERMES	HERM	Y, E	All	All	All	All	All	All	All	All	All	All	All	All
DailyDayCent	DayC	Y, E	All	All	All	All	All	Not 1x1[d]	All	All	All	Not 1x1	All	Not 1x1
APSIM-Nwheat	NWHE	Y, E	Not 1x1	All	All	All	All	All	-	-	-	-	-	-
AgroC[e]	AGRC	Y, E	All	All	All	-	-	-	All	All	All	-	-	-
MCWLA	MCLW	Y, E	All	All	All	-	-	-	-	-	-	-	-	-

[a] M_{fix} is a fixed management strategy for each crop; *Ms* indicates that sowing and fertilization dates depend on the grid cell and the year; Ms_{var}, Ms_{F50} and Ms_{F25} are the same as *Ms* but with adaptation of cultivar precocity to the cell or with a 50% and 75%, decrease in fertilization, respectively.
[b] Y is yield; E is actual evapotranspiration over the growing season for both crops; D is annual water drainage under wheat.
[c] "All" indicates that all resolutions (1x1 km, 10x10 km, 25x25 km, 50x50 km and 100x100 km) were simulated
[d] "Not 1x1" indicates that all resolutions except for 1x1 km were simulated.
[e] Data for E in AgroC are for maize only.

The simulations were done for the five resolutions (1x1, 10x10, 25x25, 50x50, and 100x100 km) with the same resolution for soil, climate, and management inputs. Among the six different management sets (M_{fix}, *Ms*, Ms_{var}, Ms_{F75}, Ms_{F50}, and Ms_{F25}), the uniform one (M_{fix}) was the same over all resolutions, while the others based on decision rules were generated at the same resolution as soil and climatic inputs. This resulted in a maximum of 30 combinations for each crop (five resolutions for each of the six management sets).

Scaling and management effects were studied on outputs averaged at the regional scale. Scaling effect was defined as the difference on the output of interest when using coarser resolution inputs in a model. Management effect was defined as the difference on the output of interest when using different management inputs in a model.

2.4. Data analysis

We quantified management and scaling effects on the regional means for each year of the 30-year simulation and for all 30 years together by model, crop and output variable. To analyze the scaling effect, we calculated the difference between the output at each resolution ($\bar{X}_{Sx}$) and those simulated at the highest resolution available ($\bar{X}_{Sr}$):

$$\Delta\bar{X}_S = \frac{\bar{X}_{Sx}-\bar{X}_{Sr}}{\bar{X}_{Sr}} \times 100 \quad [1]$$

where $\Delta\bar{X}_S$ is the difference (%) in the output at a given resolution compared to that at the reference resolution, $\bar{X}_{Sx}$ is the mean output for the region at a given resolution, and $\bar{X}_{Sr}$ is the mean output of the region at the reference resolution, which was the 1×1 km resolution, except for APSIM-Nwheat in M_{fix} and DailyDayCent in Ms_{F25} and Ms_{F75} for which it was 10×10 km. We calculated this difference due to input resolution by crop, model, and management set for each resolution, except the reference set.

To analyze the management effect, we calculated the difference between the output for each management set ($\bar{X}_{Mx}$) and those simulated for the reference set ($\bar{X}_{Mr}$):

$$\Delta\bar{X}_M = \frac{\bar{X}_{Mx}-\bar{X}_{Mr}}{\bar{X}_{Mr}} \times 100 \quad [2]$$

where $\Delta\bar{X}_M$ is the difference (%) in the output for a given management set compared to that for the reference set, $\bar{X}_{Mx}$ is the mean output for the region for a given management set x, and $\bar{X}_{Mr}$ is the mean regional output for the reference management set, which was M_{fix}, except for Expert-N, for which it was *Ms*. We calculated this difference resulting from adaptive management by crop, model, and resolution for each management set, except for the reference set.

For analyses at the annual scale, we calculated an annual scaling effect (*ASE*) and annual management effect (*AME*) for each of the 30 years, following the same logic as that for the 30-year mean (Eq. 1 and 2), but applied to the annual regional mean of each model. Again, we calculated these differences by model, crop, output, and resolution for *AME* or management set for *ASE*.

To determine if the effects of management or scaling were significant, we used a Student's *t*-test to compare each regional mean for a given output to the result of its reference (1x1km for scaling and M_{fix} for management in most cases). The comparison was done on both annual and 30-years means for each model, crop, and output.

3. Results

3.1 Simulated yield, evapotranspiration, and drainage for winter wheat and silage maize

Predictions of the regional annual yield, evapotranspiration, and drainage for the two crops differed among models for M_{fix} at 1×1 km resolution. This difference was particularly large for evapotranspiration for both crops, with regional annual medians by model ranging from 236-477 mm (235-484 mm for means) over the wheat growing season and 285-527 mm (284-523 mm for means) over the maize growing season, resulting in a maximum difference of 334 and 239 mm, respectively (Fig. 3). Regional annual median wheat yield varied less among models, from 5.8-8.0 t ha^{-1} (6.0-7.9 t ha^{-1} in mean), while median maize yield ranged from 11.3-16.2 t ha^{-1} (10.4-15.5 t ha^{-1} in mean). Median drainage varied from 356-500 mm yr^{-1} (355-497 mm yr^{-1} in mean) resulting in a maximum difference of 144 mm among the five models providing simulated drainage.

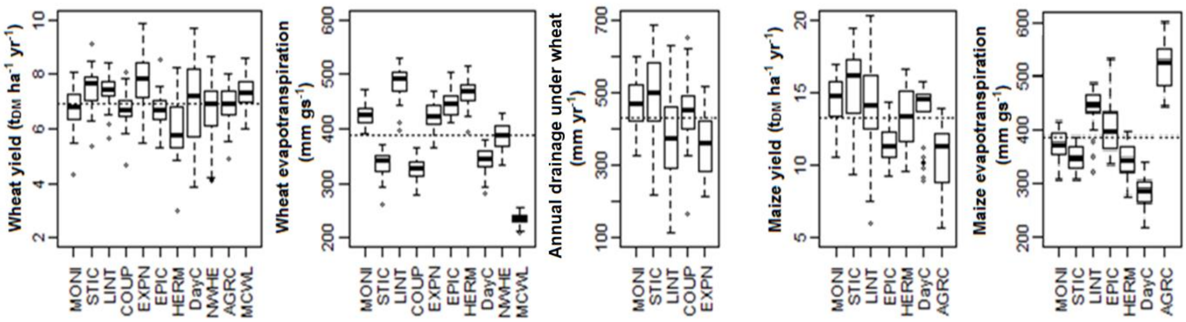

Figure 3. Distributions of the region's annual means of yield (dry matter (DM); t_{DM} ha^{-1} yr^{-1}) and evapotranspiration (growing season (gs); mm gs^{-1}) for wheat and maize, and annual mean drainage (mm yr^{-1}) under wheat over 30-year simulations by each model at its reference resolution (1 km × 1 km, except for NWHE (10 km × 10 km)) and management set (M_{fix}, except for EXPN (*Ms*)). The dotted line indicates the ensemble mean of all models for a given output. See Table 1 for model abbreviations.

Inter-annual variability also varied among the models (Fig. 3). For instance, LINTUL predicted highest inter-annual variability in maize yield, while EPIC predicted lowest variability. A similar difference was observed for wheat yield between DailyDayCent (highest) and MCWLA (lowest), and for annual drainage between MONICA (highest) and STICS (lowest).

3.2 Management effect on 30-year regional means at each resolution

We analyzed the management effect on 30-year regional means by comparing *Ms*, Ms_{var} and Ms_{F75} to M_{fix} at each resolution. Maximum management effects (in negative and positive) in yield, evapotranspiration, and drainage among models were -26% and +31%, -27% and +15%, and -12% and +1%, respectively (Table 2). For yield, these maximum management effects were similar for wheat and maize. For evapotranspiration, maximum positive differences (overestimation as compared to the reference) were slightly higher for wheat (+14%) than for maize (+4%). For maize evapotranspiration, the difference tended to be negative (underestimation as compared to the reference), whereby this trend was less consistent for wheat evapotranspiration. For drainage, the use of adaptive management sets tended to result in a negative difference (underestimation) that was the same within all resolutions, but one that was smaller than those for yield or evapotranspiration. However, the number of crop models reporting simulated drainage was much smaller as those reporting yield or evapotranspiration.

Table 2. Maximum negative and positive management effect among models ($Min(\Delta\bar{X}_M)$; $Max(\Delta\bar{X}_M)$) for the sets *Ms*, Ms_{var} and Ms_{F75} compared to M_{fix} and number of models in each level of absolute effect ($|\Delta\bar{X}_M|$) for a given output averaged over the region and all 30 years. The results are shown by crop and resolution (1 km x 1 km to 100 km x 100 km).

			Wheat					Maize					All crops
			1x1	10x10	25x25	50x50	100x100	1x1	10x10	25x25	50x50	100x100	All Res
Maximum negative and positive effect (%)	**Y**[1]	$Min(\Delta\bar{X}_M)$	-20	-18	-19	-19	-24	-18	-19	-20	-21	-26	**-26**
		$Max(\Delta\bar{X}_M)$	18	19	20	31	20	23	24	20	20	20	**31**
	E[2]	$Min(\Delta\bar{X}_M)$	-22	-23	-23	-24	-24	-21	-22	-23	-24	-27	**-27**
		$Max(\Delta\bar{X}_M)$	14	14	14	15	15	4	4	7	3	3	**15**
	D[3]	$Min(\Delta\bar{X}_M)$	-12	-12	-12	-12	-12			NA			**-12**
		$Max(\Delta\bar{X}_M)$	0	0	1	1	1						**1**
Number of models by management effect level	**Y**	$\|\Delta\bar{X}_M\| \leq 5\%$	4	5	5	5	6	0	0	0	0	0	**25**
		$5\% < \|\Delta\bar{X}_M\| \leq 10\%$	2	3	3	2	2	1	0	0	1	2	**16**
		$10\% < \|\Delta\bar{X}_M\| \leq 15\%$	1	1	1	0	0	2	4	4	3	2	**18**
		$15\% < \|\Delta\bar{X}_M\| \leq 20\%$	3	2	2	3	2	3	2	1	2	1	**21**
		$20\% < \|\Delta\bar{X}_M\| \leq 30\%$	0	0	0	0	1	1	1	2	1	2	**8**
		$30\% < \|\Delta\bar{X}_M\| \leq 40\%$	0	0	0	1	0	0	0	0	0	0	**1**
		Total	*10*	*11*	*11*	*11*	*11*	*7*	*7*	*7*	*7*	*7*	***89***
	E	$\|\Delta\bar{X}_M\| \leq 5\%$	5	5	6	6	6	2	2	3	3	3	**41**
		$5\%<\|\Delta\bar{X}_M\| \leq 10\%$	2	2	1	0	1	2	2	1	1	1	**13**
		$10\%\|\Delta\bar{X}_M\| \leq 15\%$	1	2	1	1	0	1	1	1	1	1	**10**
		$15\%<\|\Delta\bar{X}_M\| \leq 20\%$	0	0	0	1	2	1	1	1	1	1	**8**
		$20\%<\|\Delta\bar{X}_M\| \leq 30\%$	1	1	2	2	1	1	1	1	1	1	**12**
		Total	*9*	*10*	*10*	*10*	*10*	*7*	*7*	*7*	*7*	*7*	***84***
	D	$\|\Delta\bar{X}_M\| \leq 5\%$	4	4	4	4	4						**20**
		$10\%<\|\Delta\bar{X}_M\| \leq 15\%$	1	1	1	1	1			NA			**5**
		Total	*5*	*5*	*5*	*5*	*5*						***25***

[1] Y is crop yield
[2] E is evapotranspiration over the growing season
[3] D is drainage over the growing season

The response of outputs to management adaptations was model-dependent (see Table S1). For wheat, certain models had low sensitivity to management sets, such as CoupModel, Expert-N, and STICS for all outputs ($|\Delta\bar{X}_S|$ ≤ 6%) and LINTUL for yield and evapotranspiration. Other models were much more sensitive to changes in management, such as HERMES, AgroC, and DailyDayCent for crop yield, MCWLA and EPIC for evapotranspiration, and LINTUL for drainage. Overall, most predictions were similar to those with M_{fix} ($|\Delta\bar{X}_S|$ = 0-5%), although, some models predicted a large difference in the model output for certain management sets. This range of absolute difference below 5% was most common for most outputs, except for maize yield, for which the most common range of absolute difference was 10 to 15%. The regional yield for maize appeared more sensitive to differences in management than that for wheat, while the same range of differences was observed for evapotranspiration between the two crops. This higher sensitivity for maize was not related to a particular management set, since each one (Ms, Ms_{var}, Ms_{F75}) could reach the same range of absolute difference, depending on the model.

The management effect on the 30-year regional mean was similar among resolutions for a given crop and output for most models (see Table S1). Therefore, resolution did not seem to influence the difference due to management, except for APSIM-Nwheat at 50×50 km resolution for both wheat yield and evapotranspiration, and for MCWLA at a resolution of 10×10 km and coarser for wheat evapotranspiration.

Table 3. Statistical analysis by model, crop and output of the management and scaling effect. Significant difference (** p-value <0.05) were tested by Student's t-Test compared to the reference. The number of model with significant effect and the total number of model available are given at the bottom of the table.

	Management effect													
	Wheat													
	Ms			Msvar		MsF75			MsF50			MsF25		
	Yd	ET	Dr	Yd	ET	Yd	ET	Dr	Yd	ET	Dr	Yd	ET	Dr
MONI		(p,a)		(p)	(p,a)	(p)	(p,a)		(p,a)	(p)		(p,a)	(p,a)	
STIC	(p)	(p)		(p)	(p,a)	(p)	(p)		(p,a)	(p)		(p,a)	(p)	
COUP	(p)			-	-				(p)			(p,a)	-	
LINT	(p)		(p)	-	-	(p)		(p)	(p)		(p)	(p,a)	(p,a)	(p)
EPIC	(p)	(p,a)	-	(p)	(p,a)	(p)	(p,a)	-	(p,a)	(p,a)	-	(p,a)	(p,a)	-
EXPN	-	-	-			(p)			(p)			(p)	(p)	
HERM	(p,a)	(p,a)	-	(p,a)	(p,a)	(p,a)	(p,a)	-	(p)	(p,a)	-	(p,a)	(p,a)	-
DayC			-			(p,a)		-	(p,a)		-	(p,a)		-
NWHE		(p,a)	-		(p,a)		(p,a)	-		(p,a)	-		(p,a)	-
AGRC	(p,a)	-	-	(p,a)	-					-				
MCWL	(p)	(p,a)	-	(p)	(p,a)					-				
nb(p,a)	7-2	7-5	1-0	6-2	7-6	7-2	5-4	1-0	8-4	5-3	1-0	8-7	8-5	1-0
Total	10	9	4	9	8	9	9	5	9	9	5	9	8	5

	Management effect									
	Maize									
	Ms		Msvar		MsF75		MsF50		MsF25	
	Yd	ET	Yd	ET	Yd	ET	Yd	ET	Yd	ET
MONI	(p,a)	(p,a)	(p,a)	(p)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)
STIC	(p)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)
COUP					-					
LINT	(p)			(p)	(p,a)	(p)	(p,a)	(p)	(p,a)	(p)
EPIC	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)	(p,a)
EXPN					-					
HERM	(p)		(p,a)	(p,a)	(p,a)		(p,a)		(p,a)	
DayC			(p,a)				(p,a)		(p,a)	
NWHE					-					
AGRC	(p,a)	(p,a)	(p,a)	(p,a)	-	-	-	-	-	-
MCWL					-					
nb(p,a)	6-3	4-4	6-6	6-4	5-5	4-3	6-6	4-3	6-6	4-3
Total	7	7	7	7	6	6	6	6	6	6

	Scaling effect											
	Wheat								Maize			
	r10	r25		r50		r100			r50		r100	
	Dr	ET	Dr	ET	Dr	Yd	ET	Dr	Yd	ET	Yd	ET
MONI			(p)		(p)			(p)				
STIC	(p)		(p)		(p)			(p,a)				
COUP	(p)		(p)	(p)	(p)			(p,a)			-	
LINT						(a)						
EPIC	-		-		-			-	(a)		(a)	
EXPN			(p)		(p)		(p)	(p)			-	
HERM	-		-		-			-				
DayC	-	(p)	-	(p,a)	-		(p,a)	-	(a)	(p,a)		(p)
NWHE	-		-		-	(p)	(p,a)	-			-	
AGRC	-		-	-	-		-	-				
MCWL	-		-		-			-			-	
nb(p,a)	2-0	1-0	4-0	2-1	4-0	1-1	3-2	4-2	0-2	1-1	0-2	1-0
Total	5	11	5	10	5	11	10	5	7	7	7	7

"Yd" is yield, "ET" is evapotranspiration over the growing season and "Dr" is the annual drainage. "p" means that the 30-years mean is significantly different from the reference. "a" means that the annual mean is significantly different from the reference "-" means the outputs was not available for a given model. No value means that the effect was not significant. "nb(p,a)" is the number of cases for which 30-yrs and annual means were significantly different from the reference.

Management effect were significant on yield and evapotranspiration for more than half of models irrespectively of the management set used (Table 3). The effect on drainage was significant only for one of the five models that provide all three output variables, the LINTUL model. Significant effects were not linked to one management set in particular, even if they were slightly more frequent in the low fertilization management set (Ms_{F50}, Ms_{F25}) for some models.

3.3 Scaling effect on the 30-year regional means for each management set

We analyzed the scaling effect on 30-year regional means by comparing the coarser resolutions to the finest one for each of the six management sets.

Table 4. Maximum negative and positive scaling effect among models ($Min(\Delta\bar{X}_S)$; $Max(\Delta\bar{X}_S)$) for each resolution (10 km × 10 km, 25 km × 25 km, 50 km × 50 km and 100 km × 100 km) compared to the finest resolution and number of models in each level of absolute effect ($|\Delta\bar{X}_S|$) for a given output averaged over the region and all 30 years. The results are shown by crop and management set (M_{fix} to Ms_{F25}).

			Wheat						Maize						All crops
			M_{fix}	**Ms**	**Ms_{var}**	**Ms_{F75}**	**Ms_{F50}**	**Ms_{F25}**	**M_{fix}**	**Ms**	**Ms_{var}**	**Ms_{F75}**	**Ms_{F50}**	**Ms_{F25}**	**All Man**
Maximum negative and positive effect (%)	**Y**[1]	$Min(\Delta\bar{X}_S)$	-9	-8	-12	-8	-10	-15	-11	-9	-11	-8	-7	-11	**-15**
		$Max(\Delta\bar{X}_S)$	5	24	9	5	6	8	9	10	12	6	7	5	**24**
	E[2]	$Min(\Delta\bar{X}_S)$	-3	-4	-5	-3	-3	-4	-5	-3	-7	-2	-2	-3	**-7**
		$Max(\Delta\bar{X}_S)$	6	15	7	8	8	6	9	14	7	6	7	5	**15**
	D[3]	$Min(\Delta\bar{X}_S)$	-16	-15	-15	-15	-16	-16	NA						**-16**
		$Max(\Delta\bar{X}_S)$	0	0	0	0	0	0							**0**
Number of models by scaling effect level	**Y**	$|\Delta\bar{X}_S| \leq 5\%$	7	6	4	5	4	5	2	4	3	4	4	3	**51**
		$5\% < |\Delta\bar{X}_S| \leq 10\%$	4	4	3	4	4	2	4	3	2	2	2	2	**36**
		$10\% < |\Delta\bar{X}_S| \leq 15\%$	0	0	2	0	1	2	1	0	2	0	0	1	**9**
		$20\% < |\Delta\bar{X}_S| \leq 25\%$	0	1	0	0	0	0	0	0	0	0	0	0	**1**
		Total	*11*	*11*	*9*	*9*	*9*	*9*	*7*	*7*	*7*	*6*	*6*	*6*	**97**
	E	$|\Delta\bar{X}_S| \leq 5\%$	9	8	7	8	8	8	5	5	4	5	5	5	**77**
		$5\% < |\Delta\bar{X}_S| \leq 10\%$	1	1	3	1	1	1	2	1	3	1	1	1	**17**
		$10\% < |\Delta\bar{X}_S| \leq 15\%$	0	1	0	0	0	0	0	1	0	0	0	0	**2**
		Total	*10*	*10*	*10*	*9*	*9*	*9*	*7*	*7*	*7*	*6*	*6*	*6*	**96**
	D	$5\% < |\Delta\bar{X}_S| \leq 10\%$	2	3	4	3	3	3	NA						**18**
		$10\% < |\Delta\bar{X}_S| \leq 15\%$	1	1	1	1	1	1							**6**
		$15\% < |\Delta\bar{X}_S| \leq 20\%$	1	1	0	1	1	1							**5**
		Total	4	5	5	5	5	5							**29**

[1] Y is crop yield
[2] E is evapotranspiration over the growing season
[3] D is drainage over the growing season

Overall, the scaling effect on yield was in a smaller range of differences than the management effect, ranging from -15% to +24% and from -26% to +31%, respectively (Table 2 and 4). The scaling effect was weaker on evapotranspiration than on yield or drainage, with most models having an absolute difference below 5% only. Over all models, the scaling effect was both negative and positive on yield and evapotranspiration but always negative (underestimation) on drainage regardless of the model (Table 4). For the five models simulating the three output variables, evapotranspiration shows the smallest overall range with -5 to 2% while drainage and yield ranged from -16 to 0% and -10 to 3% respectively.

Certain models were more sensitive to scaling when simulating maize yield or evapotranspiration, such as STICS, and EPIC, whereas others were more sensitive when simulating wheat, such as LINTUL and DailyDayCent (see Table S2). For models predicting all three outputs, the scaling effect was higher on drainage than on yield and smallest on evapotranspiration. The scaling effect was similar across the management sets, meaning that there is no observable trend related to the management sets, regardless of the crop simulated or model used.

The significance was more frequent for management effect than for scaling on yield and evapotranspiration while it was the opposite for drainage (Table 3). The scaling effect on yield was significant only for the coarsest resolutions (100×100 km) and for one model (NWHEAT) while it was significant on three models and more resolutions (25×25 km, 50×50 km, and 100×100 km) for evapotranspiration. The scaling effect on drainage was significant for all resolutions and most models. As opposed to the management effect, the significance of scaling effect was dependent on resolution with more frequent significant effect for coarser resolutions.

3.4 Scaling and management effects at the annual scale

For the 30-year simulations, we calculated the ASE and AME on the regional means for each variable and each model. Compared to AME, ASE was much weaker on yield and evapotranspiration for both crops, particularly when excluding the 100×100 km resolution (Fig. 4). This effect was more obvious on yield than on evapotranspiration, for which the ASE and AME often remained weak, which was also the case for simulated drainage. The maximum difference due to a specific management set or resolution for a given year was also strongly model-dependent.

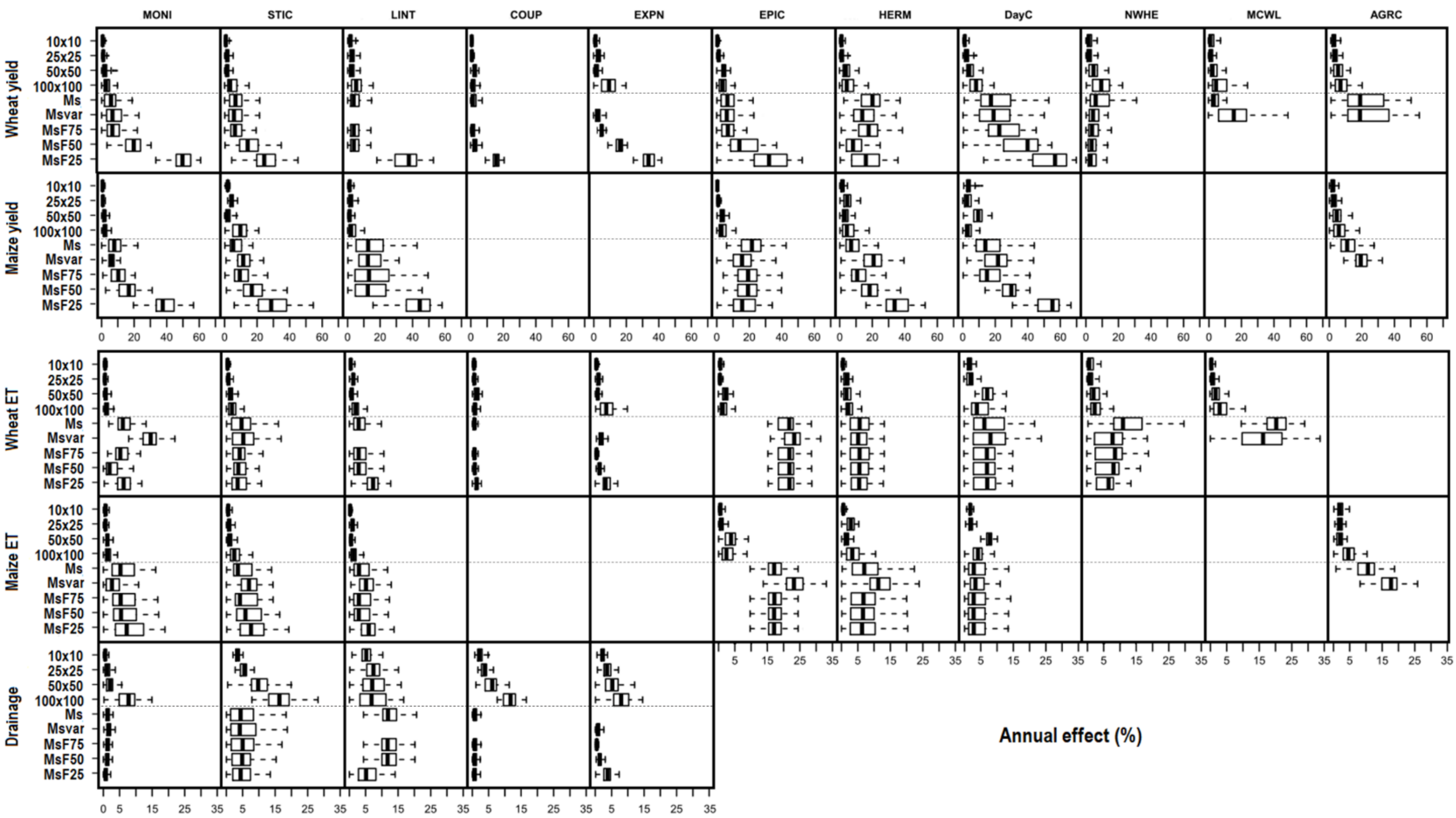

Figure 4. Distributions of annual scaling (ASE, 10x10, 25x25, 50x50 and 100x100) and management (AME, *Ms*, Ms_{var}, Ms_{F75}, Ms_{F50} and Ms_{F25}) effects on yield, evapotranspiration and drainage over 30-year simulation period compared to their respective reference, for 11 crop models (without outlier). For a given model, crop and output, each boxplot represents the average over all grid cells for each year over the

30 years, and either all management sets for the scaling effect or all resolutions for the management effect. See Table 1 for model abbreviations.

Figure 4 shows that the maximum ASE was generally small but increased with coarser resolution. For simulated wheat yield, APSIM-Nwheat had highest maximum ASE (77%) compared to the other models (<38%) at the 50×50 km resolution due to higher yield using *Ms*, while it was in the same range as those of the other models at the other resolutions. This led to a higher evapotranspiration (28%) as well on this *Ms* set and 50×50 km resolution. Apart from this set, the maximum ASE for APSIM-Nwheat at 50×50 km resolution was 19% and 6% on yield and evapotranspiration, respectively. On maize evapotranspiration, maximum ASE was highest at 25×25 km resolution for HERMES (17%) but was in the same range as those of the other models at the other resolutions (9% or lower). Generally, the models with the highest ASE were APSIM-Nwheat, DailyDayCent, HERMES, and in certain cases MCWLA, STICS, and LINTUL, depending on the output variable and the crop. In general, the ASE on yield of both crops and drainage was similar, and weakest on evapotranspiration (usually less than 10%).

The AME was generally higher on yield and evapotranspiration than ASE but had a similar range for drainage. Some models had an extremely large maximum AME, reaching 160% of the difference for a given year on the regional wheat yield for DailyDayCent and 120% on the regional maize yield for LINTUL (Fig. 4). For some models, such as CoupModel, maximum AME was around 10% only, indicating that regardless of the year, the difference due to management was low, except for Ms_{F25}, for which the maximum AME was at least 20%, regardless of the model. The AME was weaker on evapotranspiration than on yield and was even weaker on drainage. AME was similar for wheat and maize, but the difference among models was larger for wheat. This is partly because the models with the lowest AME (CoupModel and Expert-N) are available only for wheat and because maximum AME in LINTUL was higher on wheat than on maize evapotranspiration (68-71% vs. 20-25%, respectively). Drainage was less variable, with the weakest AME for the models only simulating all outputs, except for CoupModel, for which the AME was weaker on evapotranspiration. The maximum AME on drainage was 22% for LINTUL, 19% for STICS, 11% for Expert-N, 8% for CoupModel, and 4% for MONICA. No consistent trend occurred among the management sets as for evapotranspiration. Additionally, no effect of resolution on AME was observed, since the difference was the same at the five resolutions for a given crop, output variable, and model (data not shown). AME was generally low on evapotranspiration, and even lower on drainage in 90% of the situations, regardless of the model or the crop simulated, unlike regional yield, which was more sensitive to the management set.

As for the 30-years averages, significance was more frequent for management effect than for scaling (Table 3). Management influenced significantly yield and evapotranspiration under growing season for some models but not annual drainage. This result was observed for some models simulating the three output variables such as MONICA and STICS that have significant management effect for evapotranspiration or yield but not for drainage. Scaling effects were generally not significant, with some exceptions for the two coarser resolutions while management effects were often significant, especially for the two low fertilization management sets (Ms_{F50} and Ms_{F25}). Management effects were more frequently significant for maize yield and evapotranspiration than for wheat at this annual scale for most models.

4. Discussion

4.1. Management and scaling effect on the 30-year regional mean

At the multi-year scale over 30 years, the scaling and management effects were weak for most models, crops and outputs, even if significant. The scaling effect results confirm the results of previous studies on the impact of soil and climate aggregation on yield and net primary productivity (NPP) for the same study site and simulation period (Hoffmann et al., 2016b; Kuhnert et al., 2016). Further, our results indicate that varying management options over space and time in the region did not change the overall findings made when assuming constant management. Nevertheless, the scaling effect depended on the output variable, being larger for drainage than for yield or evapotranspiration when compared between the five models simulating the three output variables. The impact of the choice of the crop (winter or spring crop) on the other hand was negligible. The stronger scaling effect on drainage (observed for models providing the three outputs) and the direction of its difference was probably due to the choice of the dominant soil when moving from high to lower resolution. Lowering the resolution of soil input data resulted in an increase in the total soil water storage because deep soils were dominant in the region, which induced lower drainage. Grosz et al. (2017) also observed the scaling effect on predictions of change in soil organic carbon over time, which depend greatly on soil input data. In the same way, Coucheney et al. (2018) showed that the sensitivity to scaling was output-dependent with a greater effect of soil aggregation on soil C mineralization and N leaching than on yield and drainage for the CoupModel.

The maximum management effect tended to be higher than the maximum scaling effect, with 42 *vs.* 10 % of the cases in which differences compared to the reference were greater than 10 %, respectively. The management effect varied among models, with most 30-year regional mean outputs being slightly sensitive to management (absolute difference below 10%). This was particularly true for evapotranspiration of both crops, drainage and wheat yield, regardless of the input resolution. The stronger effect on yield could be partly due to the use of percentage to quantify the effect. Since average yields are much lower than evapotranspiration and drainage, a small variation lead to a higher percentage for this output. However, for the scaling effect the effect was strongest on drainage. The management effect tended to be higher on maize than on winter wheat yield for most models, suggesting a greater impact of management on spring crops than on winter crops. This result seems consistent with the shorter growing season of spring crops, leaving less time to compensate a late sowing for instance. The hypothesis of a higher sensitivity of spring crops should be tested with other crops such as sunflower or soybean. For some models (2-4 models), different representation of management changed the 30-year regional mean substantially (by more than 15% for yield and for evapotranspiration depending on the resolution and crop), indicating the need to carefully choose how to represent management in these crop models to obtain relevant multi-year regional means. Contrary, management choices seemed less important for the 30-year regional drainage, (showing less than 13 % difference in all management sets).

4.2. Stronger effects at the annual scale

The same trend occurred at the annual scale as for the 30-year regional mean: the management effect was usually higher than the scaling effect, with large differences among models. The management effect as well as the scaling effect on the regional mean were stronger for certain years than for others. This indicates that the choices made to represent management are more important when studies focus on annual regional outputs than on multi-year average regional outputs. This importance varied among models and, depending on the model, the cultivation operation considered. Hereby, it is crucial to ensure that the chosen model is able to predict effects of a given management strategy, such as sowing date, to accurately predict variability in the outputs caused by the management changes. If the management strategy has a substantial effect on the output variable of interest, the uncertainty due to the choice of management option in the simulation should be estimated.

Since the years with large effects on management options or scaling differed among models, it is difficult to identify which characteristics of the years that interact with the models to generate the more or less strong effects. No effect of climate characteristics such as a dry or hot year effects was found in the analyses. The effect were strongly model-dependent, the same year predictions being sensitive to scaling or management effect for some models but not for others. No generic characteristics of the input data could be identified; the effect being probably due to a model-soil-climate interaction. This difference between crop model outputs behavior is probably partly due to model structures as well as their parametrization, their the relative contribution being unclear. Hereby, sensitivity analysis performed in individual studies of each model could be helpful to understand model behavior and to determine characteristic input-output relationships (Specka et al., 2015; Varella et al., 2012). It could then clarify the major factors behind model differences with respect to the occurrence of strong effects of management strategies in specific years.

4.3. Representation of management strategies in large-scale studies

We used decision rules to generate management options based on climatic conditions. We then compared simulations based on these management options with those of uniform and fixed sowing, harvest, and fertilization dates over one region over multiple years. In general, uniform sowing, harvest, and fertilization dates as well the use of a single cultivar are an unrealistic representation of common management at the regional scale. Folberth et al. (2016) showed that in model-based global scale assessments, absolute yield levels depend on the parameterization and distribution of crop cultivars. However, it is still commonly applied in large-scale modelling studies since real data are often scarce (Faivre et al., 2004). The advantage of using decision rules is that it provide a management, which is consistent with local climate and soil as compared to fixed assumptions. These can also be used to simulate changes in management over time due to climate change (Senthilkumar et al., 2015). One limitation is that the same decision rules are used for all grid cells, while different farmers apply different rules for crop management (Maton et al., 2005) depending on their social, economic, and pedoclimatic conditions. Decisions rules based on an optimal strategy according to climatic conditions could lead to overestimated yields. Moreover, not taking into account soil characteristics could also lead to unrealistic management in some cases. Since the purpose of this study was to evaluate if management choices had an impact on regional output variables, these concerns were not of critical importance. To get more realistic data on management at large scale, remote sensing could add useful information on crop type (Griffiths et al., 2019), sowing and harvest dates, or irrigation schedules (Battude et al., 2017). Here, we analyzed the potential impact of choosing a variable management to predict the difference in crop model outputs compared to a reference based on a spatially uniform management fixed in time. The access and use of observed management data for the entire region to validate the relevance and accuracy of the decision rules, would improve assessments of the role and effect of management input data and resolution for simulations at regional scale. It could be relevant to include other cultivation operations, such as soil tillage or irrigation, depending on the outputs of interest. For instance, irrigation is important when water balance is the focus of the simulation study, particularly in southern Europe.

5. Conclusion

In our regional-scale study, we showed that the management effect was generally stronger than the scaling effect. The strength of the effects depended on the crop model and the output variable of interest, with some models and output variables being much more sensitive to management options than others. Scaling and management effects were also stronger when evaluated on individual years than on the 30-year mean, for which these effects were usually weak. The effects varied both between models and among years. Strong impacts occurred but not necessarily during the same years for all models, which indicates a need for further analysis with respect to each model to explain these effects in depth. Additionally, the findings of this study might be different in other conditions and therefore need to be confirmed with respect to a different region with contrasting soil and climate conditions.

Acknowledgments

This work was supported by the FACCE MACSUR knowledge hub (http://macsur.eu). JC, HR, EC and JEB thank the INRA ACCAF metaprogramme for funding. FT and RPR were supported by FACCE MACSUR (3200009600) through the Finnish Ministry of Agriculture and Forestry (MMM). HE, EL and AV were supported by The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (220-2007-1218) and by the strategic funding 'Soil-Water-Landscape' from the faculty of Natural Resources and Agricultural Sciences (Swedish University of Agricultural Sciences) and thank professor P-E Jansson (Royal Institute of Technology, Stockholm) for support. ET was funded by the Royal Society of New Zealand and the Climate Change Impacts and Implications for New Zealand project (CCII) financed by the Ministry of Business, Innovation and Employment (MBIE). FE, TG and HH acknowledge support by the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE), (2851ERA01J). KCK, CN and XS acknowledge FACCE MACSUR (2812ERA147). MK and JY thank for the funding by the UK BBSRC (BB/N004922/1) and the MAXWELL HPC team of the University of Aberdeen for providing equipment and support through the German Federal Ministry of Food and Agriculture for the DailyDayCent simulations. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Asseng, S., van Keulen, H., Stol, W., 2000. Performance and application of the APSIM Nwheat model in the Netherlands. Eur. J. Agron. 12, 37–54. https://doi.org/10.1016/S1161-0301(99)00044-1

Battude, M., Al Bitar, A., Brut, A., Tallec, T., Huc, M., Cros, J., Weber, J.J., Lhuissier, L., Simonneaux, V., Demarez, V., 2017. Modeling water needs and total irrigation depths of maize crop in the south west of France using high spatial and temporal resolution satellite imagery. Agric. Water Manag. 189, 123–136. https://doi.org/10.1016/j.agwat.2017.04.018

Bergez, J.-E., Raynal, H., Launay, M., Beaudoin, N., Casellas, E., Caubel, J., Chabrier, P., Coucheney, E., Dury, J., García de Cortázar-Atauri, I., Justes, É., Mary, B., Ripoche, D., Ruget, F., 2014. Evolution of the STICS crop model to tackle new environmental issues: New formalisms and integration in the modelling and simulation platform RECORD. Environ. Model. Softw. 62, 370–384. https://doi.org/10.1016/j.envsoft.2014.07.010

Bonelli, L.E., Monzon, J.P., Cerrudo, A., Rizzalli, R.H., Andrade, F.H., 2016. Maize grain yield components and source-sink relationship as affected by the delay in sowing date. F. Crop. Res. 198, 215–225. https://doi.org/10.1016/j.fcr.2016.09.003

Brisson, N., Gary, C., Justes, É., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussiere, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillère, J.P., Hénault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model STICS. Eur. J. Agron. 18, 309–332. https://doi.org/10.1016/S1161-0301(02)00110-7

Conrad, Y., 2009. Modelling of nitrogen leaching under a complex winter wheat and red clover crop rotation in a drained agricultural field. Phys. Chem. Earth, Parts A/B/C 34, 530–540. https://doi.org/10.1016/J.PCE.2008.08.003

Coucheney, E., Eckersten, H., Hoffmann, H., Jansson, P.E., Gaiser, T., Ewert, F., Lewan, E., 2018. Key functional soil types explain data aggregation effects on simulated yield, soil carbon, drainage and nitrogen leaching at a regional scale. Geoderma 318, 167–181. https://doi.org/10.1016/j.geoderma.2017.11.025

De Wit, A., Duveiller, G., Defourny, P., 2012. Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations. Agric. For. Meteorol. 164, 39–52. https://doi.org/10.1016/j.agrformet.2012.04.011

Del Grosso, S.J., Parton, W.J., Mosier, A.R., Walsh, M.K., Ojima, D.S., Thornton, P.E., 2006. DAYCENT National-Scale Simulations of Nitrous Oxide Emissions from Cropped Soils in the United States. J. Environ. Qual. 35, 1451. https://doi.org/10.2134/jeq2005.0160

Ewert, F., van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepkina, I., Andersen, E., 2011. Scale changes and model linking methods for integrated assessment of agri-environmental systems. Agric. Ecosyst. Environ. 142, 6–17. https://doi.org/10.1016/j.agee.2011.05.016

Faivre, R., Leenhardt, D., Voltz, M., Benoît, M., Papy, F., Dedieu, G., Wallach, D., 2004. Spatialising crop models. Agronomie 24, 205–217. https://doi.org/10.1051/agro

Folberth, C., Elliott, J., Müller, C., Balkovič, J., Izaurralde, R.C., Jones, C.D., Khabarov, N., Liu, W., Reddy, A., Schmid, E., Skalský, R., Yang, H., 2016. Uncertainties in global crop model frameworks : effects of cultivar distribution , crop management and soil handling on crop yield estimates 1–30. https://doi.org/10.5194/bg-2016-527

Gaiser, T., Abdel-Razek, M., Bakara, H., 2009. Modeling carbon sequestration under zero-tillage at the regional scale. II. The influence of crop rotation and soil type. Ecol. Modell. 220, 3372–3379. https://doi.org/10.1016/j.ecolmodel.2009.08.001

Gaiser, T., Judex, M., Hiepe, C., Kuhn, A., 2010. Regional simulation of maize production in tropical savanna fallow systems as affected by fallow availability. Agric. Syst. 103, 656–665. https://doi.org/10.1016/j.agsy.2010.08.004

Gaiser, T., Perkons, U., Küpper, P.M., Kautz, T., Uteau-Puschmann, D., Ewert, F., Enders, A., Krauss, G., 2013. Modeling biopore effects on root growth and biomass production on soils with pronounced sub-soil clay accumulation. Ecol. Modell. 256, 6–15. https://doi.org/10.1016/J.ECOLMODEL.2013.02.016

Gaiser, T., Stahr, K., Billen, N., Mohammad, M.A.R., 2008. Modeling carbon sequestration under zero tillage at the regional scale. I. The effect of soil erosion. Ecol. Modell. 218, 110–120. https://doi.org/10.1016/j.ecolmodel.2008.06.025

Griffiths, P., Nendel, C., Hostert, P., 2019. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. Remote Sens. Environ. 220, 135–151. https://doi.org/10.1016/j.rse.2018.10.031

Grosz, B., Dechow, R., Gebbert, S., Hoffmann, H., Zhao, G., Constantin, J., Raynal, H., Wallach, D., Coucheney, E., Lewan, E., Eckersten, H., Specka, X., Kersebaum, K.C., Nendel, C., Kuhnert, M., Yeluripati, J.B., Haas, E., Teixeira, E.I., Bindi, M., Trombi, G., Moriondo, M., Doro, L., Roggero, P.P., Zhao, Z., Wang, E., Tao, F., Rötter, R.P., Kassie, B., Cammarano, D., Asseng, S., Weihermüller, L., Siebert, S., Gaiser, T., Ewert, F., 2017. The implication of input data aggregation on up-scaling soil organic carbon changes. Environ. Model. Softw. 96, 361–377. https://doi.org/10.1016/j.envsoft.2017.06.046

Herbst, M., Hellebrand, H.J., Bauer, J., Huisman, J.A., Šimůnek, J., Weihermüller, L., Graf, A., Vanderborght, J., Vereecken, H., 2008. Multiyear heterotrophic soil respiration: Evaluation of a coupled CO_2transport and carbon turnover model. Ecol. Modell. 214, 271–283. https://doi.org/10.1016/j.ecolmodel.2008.02.007

Hoffmann, H., Enders, A., Siebert, S., Gaiser, T., Ewert, F., 2016a. Climate and soil input data aggregation effects in crop models. Havard Database V3. https://doi.org/https://doi.org/10.7910/DVN/C0J5BB

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E., Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B.T., Kersebaum, K.C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., Nendel, C., Priesack, E., Raynal, H., Roggero, P.P., Rötter, R.P., Siebert, S., Specka, X., Tao, F., Teixeira, E.I., Trombi, G., Wallach, D., Weihermüller, L., Yeluripati, J.B., Ewert, F., 2016b. Impact of Spatial Soil and Climate Input Data Aggregation on Regional Yield Simulations. PLoS One 11, e0151782. https://doi.org/10.1371/journal.pone.0151782

Hoffmann, H., Zhao, G., Van Bussel, L.G.J., Enders, A., Specka, X., Sosa, C., Yeluripati, J.B., Tao, F., Constantin, J., Raynal, H., Teixeira, E.I., Grosz, B., Doro, L., Zhao, Z., Wang, E., Nendel, C., Kersebaum, K.C., Haas, E., Kiese, R., Klatt, S., Eckersten, H., Vanuytrecht, E., Kuhnert, M., Lewan, E., Rötter, R.P., Roggero, P.P., Wallach, D., Cammarano, D., Asseng, S., Krauss, G., Siebert, S., Gaiser, T., Ewert, F., 2015. Variability of effects of spatial climate data aggregation on regional yield simulation by crop models. Clim. Res. 69, 53–69. https://doi.org/10.3354/cr01326

Huang, J., Tian, L., Liang, S., Ma, H., Becker-Reshef, I., Huang, Y., Su, W., Zhang, X., Zhu, D., Wu, W., 2015. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. Agric. For. Meteorol. 204, 106–121. https://doi.org/10.1016/j.agrformet.2015.02.001

Hutchings, N.J., Reinds, G.J., Leip, A., Wattenbach, M., Bienkowski, J.F., Dalgaard, T., Dragosits, U., Drouet, J.L., Durand, P., Maury, O., De Vries, W., 2012. A model for simulating the timelines of field operations at a European scale for use in complex dynamic models. Biogeosciences 9, 4487–4496. https://doi.org/10.5194/bg-9-4487-2012

Jansson, P.-E., 2012. Coupmodel: Model use, calibration, and validation. Trans. ASABE 55, 1335–1344.

Kersebaum, K.C., 2007. Modelling nitrogen dynamics in soil–crop systems with HERMES, in: Modelling Water and Nutrient Dynamics in Soil–crop Systems. Springer Netherlands, Dordrecht, pp. 147–160.

https://doi.org/10.1007/978-1-4020-4479-3_11

Kersebaum, K.C., Nendel, C., 2014. Site-specific impacts of climate change on wheat production across regions of Germany using different CO2response functions. Eur. J. Agron. 52, 22–32. https://doi.org/10.1016/j.eja.2013.04.005

Klosterhalfen, A., Herbst, M., Weihermüller, L., Graf, A., Schmidt, M., Stadler, A., Schneider, K., Subke, J.A., Huisman, J.A., Vereecken, H., 2017. Multi-site calibration and validation of a net ecosystem carbon exchange model for croplands. Ecol. Modell. 363, 137–156. https://doi.org/10.1016/j.ecolmodel.2017.07.028

Kollas, C., Kersebaum, K.C., Nendel, C., Manevski, K., Müller, C., Palosuo, T., Armas-Herrera, C.M., Beaudoin, N., Bindi, M., Charfeddine, M., Conradt, T., Constantin, J., Eitzinger, J., Ewert, F., Ferrise, R., Gaiser, T., García de Cortázar-Atauri, I., Giglio, L., Hlavinka, P., Hoffmann, H., Hoffmann, M.P., Launay, M., Manderscheid, R., Mary, B., Mirschel, W., Moriondo, M., Olesen, J.E., Öztürk, I., Pacholski, A., Ripoche-Wachter, D., Roggero, P.P., Roncossek, S., Rötter, R.P., Ruget, F., Sharif, B., Trnka, M., Ventrella, D., Waha, K., Wegehenkel, M., Weigel, H.-J., Wu, L., 2015. Crop rotation modelling—A European model intercomparison. Eur. J. Agron. 70, 98–111. https://doi.org/10.1016/j.eja.2015.06.007

Kuhnert, M., Yeluripati, J.B., Smith, P., Hoffmann, H., van Oijen, M., Constantin, J., Coucheney, E., Dechow, R., Eckersten, H., Gaiser, T., Grosz, B., Haas, E., Kersebaum, K.C., Kiese, R., Klatt, S., Lewan, E., Nendel, C., Raynal, H., Sosa, C., Specka, X., Teixeira, E.I., Wang, E., Weihermüller, L., Zhao, G., Zhao, Z., Ogle, S., Ewert, F., 2016. Impact analysis of climate data aggregation at different spatial scales on simulated net primary productivity for croplands. Eur. J. Agron. 88, 41–52. https://doi.org/10.1016/j.eja.2016.06.005

Leenhardt, D., Lemaire, P., 2002. Estimating the spatial and temporal distribution of sowing dates for regional water management. Agric. Water Manag. 55, 37–52. https://doi.org/10.1016/S0378-3774(01)00183-4

Maton, L., Leenhardt, D., Goulard, M., Bergez, J.-E., 2005. Assessing the irrigation strategies over a wide geographical area from structural data about farming systems. Agric. Syst. 86, 293–311. https://doi.org/10.1016/j.agsy.2004.09.010

Mo, X., Liu, S., Lin, Z., Xu, Y., Xiang, Y., McVicar, T.R., 2005. Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain. Ecol. Modell. 183, 301–322. https://doi.org/10.1016/j.ecolmodel.2004.07.032

Nendel, C., 2009. Evaluation of Best Management Practices for N fertilisation in regional field vegetable production with a small-scale simulation model. Eur. J. Agron. 30, 110–118. https://doi.org/10.1016/j.eja.2008.08.003

Nendel, C., Berg, M., Kersebaum, K.C., Mirschel, W., Specka, X., Wegehenkel, M., Wenkel, K.O., Wieland, R., 2011. The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. Ecol. Modell. 222, 1614–1625. https://doi.org/10.1016/J.ECOLMODEL.2011.02.018

Nendel, C., Wieland, R., Mirschel, W., Specka, X., Guddat, C., Kersebaum, K.C., 2013. Simulating regional winter wheat yields using input data of different spatial resolution. F. Crop. Res. 145, 67–77. https://doi.org/10.1016/j.fcr.2013.02.014

Noory, H., van der Zee, S.E.A.T.M., Liaghat, A.-M., Parsinejad, M., van Dam, J.C., 2011. Distributed agro-hydrological modeling with SWAP to improve water and salt management of the Voshmgir Irrigation and Drainage Network in Northern Iran. Agric. Water Manag. 98, 1062–1070. https://doi.org/10.1016/j.agwat.2011.01.013

Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Global Biogeochem. Cycles 24, 1–24. https://doi.org/10.1029/2008GB003435

Priesack, E., Gayler, S., Hartmann, H.P., 2006. The impact of crop growth sub-model choice on simulated water and nitrogen balances. Nutr. Cycl. Agroecosystems 75, 1–13. https://doi.org/10.1007/s10705-006-9006-1

Robert, M., Thomas, A., Sekhar, M., Raynal, H., Casellas, E., Casel, P., Chabrier, P., Joannon, A., Bergez, J.-E., 2018. A dynamic model for water management at the farm level integrating strategic, tactical and operational decisions. Environ. Model. Softw. 100, 123–135. https://doi.org/10.1016/j.envsoft.2017.11.013

Sacks, W.J., Deryng, D., Foley, J.A., Ramankutty, N., 2010. Crop planting dates: An analysis of global patterns. Glob. Ecol. Biogeogr. 19, 607–620. https://doi.org/10.1111/j.1466-8238.2010.00551.x

Savin, I., Boogaard, H., Diepen, C. Van, Ham, H. Van Der, 2007. Climatically Optimal Planting Dates. JRC Sci. Tech. Reports 58.

Senthilkumar, K., Bergez, J.-E., Leenhardt, D., 2015. Can farmers use maize earliness choice and sowing dates to cope with future water scarcity? A modelling approach applied to south-western France. Agric. Water Manag. 152, 125–134. https://doi.org/10.1016/j.agwat.2015.01.004

Specka, X., Nendel, C., Wieland, R., 2015. Analysing the parameter sensitivity of the agro-ecosystem model MONICA for different crops. Eur. J. Agron. 71, 73–87. https://doi.org/10.1016/j.eja.2015.08.004

Tao, F., Yokozawa, M., Zhang, Z., 2009. Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis. Agric. For. Meteorol. 149, 831–850. https://doi.org/10.1016/j.agrformet.2008.11.004

Tao, F., Zhang, Z., 2013. Climate change, wheat productivity and water use in the North China Plain: A new super-ensemble-based probabilistic projection. Agric. For. Meteorol. 170, 146–165. https://doi.org/10.1016/J.AGRFORMET.2011.10.003

Therond, O., Sibertin-blanc, C., Lardy, R., Gaudou, B., Sauvage, S., Taillandier, P., Vavasseur, M., Mazzega, P., Balestrat, M., Ong, Y., Louail, T., Mayor, E., Bai, V., 2014. Integrated modelling of social-ecological systems : The MAELIA high-resolution multi-agent platform to deal with water scarcity problems, in: 7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA. June, p. 8.

Tornquist, C.G., Gassman, P.W., Mielniczuk, J., Giasson, E., Campbell, T., 2009. Spatially explicit simulations of soil C dynamics in Southern Brazil: Integrating century and GIS with i_Century. Geoderma 150, 404–414. https://doi.org/10.1016/j.geoderma.2009.03.001

van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance-A review. F. Crop. Res. 143, 4–17. https://doi.org/10.1016/j.fcr.2012.09.009

Varella, H., Buis, S., Launay, M., Guérif, M., 2012. Global sensitivity analysis for choosing the main soil parameters of a crop model to be determined. Agric. Sci. 3, 949–961.

Williams, J.R., 1995. The EPIC model. Computer models of watershed hydrology. Water resources publications, Highland Ranch CO.

Williams, J.R., Renard, K.G., Dyke, P.T., 1983. EPIC: a new method for assessing erosion's effect on soil productivity. J. Soil Water Conserv. 38, 381–383.

Yeluripati, J.B., van Oijen, M., Wattenbach, M., Neftel, A., Ammann, A., Parton, W.J., Smith, P., 2009. Bayesian calibration as a tool for initialising the carbon pools of dynamic soil models. Soil Biol. Biochem. 41, 2579–2583. https://doi.org/10.1016/J.SOILBIO.2009.08.021

Zhao, G., Hoffmann, H., Van Bussel, L.G.J., Enders, A., Specka, X., Sosa, C., Yeluripati, J.B., Tao, F., Constantin, J., Raynal, H., Teixeira, E.I., Grosz, B., Doro, L., Zhao, Z., Nendel, C., Kiese, R., Eckersten, H., Haas, E., Vanuytrecht, E., Wang, E., Kuhnert, M., Trombi, G., Moriondo, M., Bindi, M., Lewan, E., Bach, M., Kersebaum, K.C., Rötter, R.P., Roggero, P.P., Wallach, D., Cammarano, D., Asseng, S., Krauss, G., Siebert, S., Gaiser, T., Ewert, F., 2015a. Effect of weather data aggregation on regional crop simulation for different crops, production conditions, and response variables. Clim. Res. 65, 141–157.

https://doi.org/10.3354/cr01301

Zhao, G., Siebert, S., Enders, A., Rezaei, E.E., Yan, C., Ewert, F., 2015b. Demand for multi-scale weather data for regional crop modeling. Agric. For. Meteorol. 200, 156–171. https://doi.org/10.1016/j.agrformet.2014.09.026

. Supplementary Material

S1. Minimum and maximum values for management effect for the sets *Ms*, Ms_{var} and Ms_{F75} compared to *Mfix* for a given output, model, crop and resolution (% difference compared to the reference management set). Brackets contain minimum and maximum differences; a single value indicates that the minimum equals the maximum. Y is crop yield, E is evapotranspiration over the growing season and D is drainage over the growing season. See Table 1 for model abbreviations.

		Wheat					**Maize**					**All crops**
		1x1	**10x10**	**25x25**	**50x50**	**100x100**	**1x1**	**10x10**	**25x25**	**50x50**	**100x100**	**All Res**
Y	**MONI**	[-6;0]	[-5;0]	[-6;0]	[-6;-1]	[-6;0]	[-11;-6]	[-11;-6]	[-11;-6]	[-10;-6]	[-9;-6]	**[-11;0]**
	STIC	[-5;6]	[-5;5]	[-5;5]	[-4;6]	[-6;4]	[-12;-7]	[-12;-7]	[-12;-7]	[-12;-6]	[-11;-4]	**[-12;6]**
	LINT	[4;5]	4	4	4	4	[-3;8]	[-3;10]	[-2;11]	[-2;12]	[-2;14]	**[-3;14]**
	COUP	[1;2]	[1;2]	[1;2]	[1;2]	2	NA	NA	NA	NA	NA	**[1;2]**
	EXPN	-3	[-5;-2]	[-5;-2]	[-5;-3]	[-5;-4]	NA	NA	NA	NA	NA	**[-5;-2]**
	EPIC	[-5;4]	[-4;4]	[-4;5]	[-4;5]	[-4;3]	[18;23]	[18;24]	[15;20]	[14;20]	[11;20]	**[-5;24]**
	HERM	[16;18]	[16;19]	[15;20]	[14;19]	[18;20]	[-18;-5]	[-19;-6]	[-20;1]	[-21;-6]	[-23;-6]	**[-23;20]**
	DayC	[-16;-2]	[-15;-1]	[-14;0]	[-15;-1]	[-15;1]	[-15;-4]	[-11;-3]	[-11;-2]	[-13;-4]	[-9;0]	**[-16;1]**
	NWHE	NA	[4;5]	[4;5]	[2;31]	[-1;3]	NA	NA	NA	NA	NA	**[-1;31]**
	AGRC	[-20;-19]	-18	[-19;-18]	[-19;-17]	[-24;-20]	[-17;12]	[-18;9]	[-19;11]	[-20;12]	[-26;11]	**[-26;12]**
	MCWL	-12	[-7;0]	[-5;4]	[-2;3]	[2;3]	NA	NA	NA	NA	NA	**[-12;4]**
E	**MONI**	[6;14]	[6;14]	[6;14]	[6;15]	[6;15]	[-6;2]	[-5;2]	[-5;3]	[-4;3]	[-4;3]	**[-6;15]**
	STIC	[2;4]	[3;4]	[3;4]	3	[2;2]	[-7;-6]	[-7;-5]	[-7;-5]	[-7;-4]	[-7;-4]	**[-7;4]**
	LINT	0	0	0	0	0	[-2;1]	[-2;1]	[-2;2]	[-2;2]	[-2;2]	**[-2;2]**
	COUP	-1	0	-1	-1	[-1;0]	NA	NA	NA	NA	NA	**[-1;0]**
	EXPN	0	[-1;0]	[-1;0]	[-2;0]	[-3;0]	NA	NA	NA	NA	NA	**[-3;0]**
	EPIC	-22	[-23;-22]	[-23;-21]	[-24;-21]	[-24;-22]	[-21;-16]	[-22;-16]	[-23;-17]	[-24;-18]	[-27;-20]	**[-27;-16]**
	HERM	[-6;-4]	[-5;-4]	[-5;-4]	-4	-3	[-11;-4]	[-10;-3]	[-10;7]	[-11;-2]	[-12;-2]	**[-12;7]**
	DayC	[-1;3]	[0;3]	[1;3]	[2;3]	[3;4]	[1;4]	[2;4]	[2;4]	[0;2]	[-1;3]	**[-1;4]**
	NWHE	NA	9	10	[9;26]	[8;10]	NA	NA	NA	NA	NA	**[0;26]**
	AGRC	NA	NA	NA	NA	NA	[-17;-11]	[-16;-12]	[-16;-9]	[-16;-9]	[-19;-9]	**[-19;-9]**
	MCWL	[-6;-6]	[0;14]	[15;21]	[15;20]	[16;18]	NA	NA	NA	NA	NA	**[-6;21]**
D	**MONI**	[-2;-1]	[-2;-1]	[-2;-1]	[-2;-1]	-2						**[-2;-1]**
	STIC	[-5;-4]	[-5;-4]	[-5;-4]	[-5;-4]	-4						**[-5;-4]**
	LINT	[-12;-11]	-12	-12	-12	-12			NA			**[-12;-11]**
	COUP	-1	-1	-1	-1	-1						**-1**
	EXPN	0	[-2;0]	[0;1]	[0;1]	[0;1]						**[-2;1]**

S2. Minimum and maximum values for scaling effect between all resolutions compared to the reference resolution (1 km × 1 km in most cases) for a given output, model, management set and crop (% difference compared to the reference resolution). Brackets contain minimum and maximum values; a single value indicates that the minimum equals the maximum. Y is crop yield, E is evapotranspiration over the growing season and D is drainage over the growing season. See Table 1 for model abbreviations.

		Wheat						**Maize**						**All crops**
		M_{fix}	*Ms*	Ms_{var}	Ms_{F75}	Ms_{F50}	Ms_{F25}	M_{fix}	*Ms*	Ms_{var}	Ms_{F75}	Ms_{F50}	Ms_{F25}	**All Mx**
Y	**MONI**	[0;1]	[0;1]	[0;2]	[0;1]	[0;2]	[0;2]	[-1;0]	[0;1]	0	[0;1]	[0;2]	[0;2]	**[-1;2]**
	STIC	[-4;1]	[-5;1]	[-6;-1]	[-5;1]	[-4;2]	[-3;2]	[-11;-2]	[-9;-1]	[-11;-2]	[-8;-1]	[-7;-1]	[-7;-1]	**[-11;2]**
	LINT	[-5;-2]	[-6;-2]	NA	[-5;-2]	[-5;-2]	[-5;-2]	[-6;-3]	[-1;0]	[-5;-2]	[-1;0]	[-1;0]	[-1;0]	**[-6;0]**
	COUP	[-1;2]	[-1;3]	NA	[-1;3]	[-1;3]	[-1;3]	NA	NA	NA	NA	NA	NA	**[-1;3]**
	EXPN	NA	[-8;-1]	[-10;-2]	[-8;-1]	[-8;-1]	[-8;0]	NA	NA	NA	NA	NA	NA	**[-10;0]**
	EPIC	[-1;4]	[0;5]	[-1;4]	[-1;5]	[-2;4]	[-3;5]	[-1;6]	[-1;3]	[-1;3]	[-1;3]	[-1;3]	[-1;3]	**[-3;6]**
	HERM	[-1;3]	[-1;4]	[-3;0]	[0;4]	[0;6]	[0;8]	[-3;3]	[-4;4]	[-9;0]	[-4;2]	[-5;3]	[-5;5]	**[-9;8]**
	DayC	[-9;0]	[-5;3]	[-6;2]	[-8;2]	[-10;3]	[-15;0]	[-5;9]	[-1;10]	[1;12]	[-4;6]	[-2;7]	[-11;2]	**[-15;12]**
	NWHE	[-7;-1]	[-8;24]	[-12;1]	[-8;0]	[-9;-1]	[-12;-5]	NA	NA	NA	NA	NA	NA	**[-12;24]**
	AGRC	[1;5]	[3;7]	[-2;3]	NA	NA	NA	[2;6]	[0;6]	[-10;2]	NA	NA	NA	**[-10;7]**
	MCWL	[-3;0]	[-4;0]	[2;9]	NA	NA	NA	NA	NA	NA	NA	NA	NA	**[-4;9]**
E	**MONI**	0	[0;1]	[0;1]	[0;1]	[0;1]	[0;1]	[-1;0]	[0;1]	0	1	1	[1;2]	**[-1;2]**
	STIC	[0;1]	[0;1]	[-2;0]	[0;1]	[0;2]	[0;2]	[-4;-1]	[-2;1]	[-4;0]	[-2;1]	[-1;1]	[-1;1]	**[-4;2]**
	LINT	[-2;0]	[-2;1]	NA	[-2;0]	[-2;0]	[-1;0]	[-2;0]	[-1;0]	[-2;0]	[-1;0]	[-1;0]	[0;1]	**[-2;1]**
	COUP	[-1;1]	[0;1]	NA	[0;1]	[0;1]	[0;1]	NA	NA	NA	NA	NA	NA	**[-1;1]**
	EXPN	NA	[-3;0]	[-5;-1]	[-3;0]	[-3;0]	[-2;1]	NA	NA	NA	NA	NA	NA	**[-5;1]**
	EPIC	[-1;2]	[-1;3]	[-3;0]	[-1;3]	[-1;3]	[-1;3]	[0;6]	[0;4]	[-2;1]	[0;4]	[0;4]	[0;4]	**[-3;6]**
	HERM	[-1;1]	[1;2]	[0;1]	[1;2]	[1;2]	[0;2]	[-4;1]	[-2;14]	[-6;2]	[-2;3]	[-2;3]	[-3;3]	**[-6;14]**
	DayC	[1;6]	[2;8]	[1;7]	[2;8]	[2;8]	[0;6]	[1;9]	[2;7]	[2;7]	[0;6]	[2;7]	[0;5]	**[0;9]**
	NWHE	[-2;-1]	[-2;15]	[-4;0]	[-2;0]	[-3;-1]	[-4;-2]	NA	NA	NA	NA	NA	NA	**[-4;15]**
	AGRC	NA	NA	NA	NA	NA	NA	[-5;-2]	[-3;0]	[-7;-1]	NA	NA	NA	**[-7;0]**
	MCWL	[-1;1]	[-4;0]	[0;2]	NA	NA	NA	NA	NA	NA	NA	NA	NA	**[-4;2]**
D	**MONI**	[-6;0]	[-7;0]	[-7;0]	[-7;0]	[-7;0]	[-7;0]							**[-7;0]**
	STIC	[-16;-3]	[-15;-3]	[-15;-3]	[-15;-3]	[-16;-3]	[-16;-3]							**[-16;-3]**
	LINT	[-6;-4]	[-6;-4]	NA	[-7;-5]	[-7;-5]	[-7;-5]			NA				**[-7;-4]**
	COUP	[-11;-2]	[-12;-2]	NA	[-12;-2]	[-12;-2]	[-12;-2]							**[-12;-2]**
	EXPN	NA	[-8;-2]	[-7;-2]	[-8;-2]	[-8;-2]	[-9;-2]							**[-9;-2]**